

STAFF SUMMARY SHEET

	TO	ACTION	SIGNATURE (Surname), GRADE AND DATE		TO	ACTION	SIGNATURE (Surname), GRADE AND DATE
1	DFPY	sig	(sig) <i>Leonard Kahn</i> 0-6 (print) 1 ^c 13 Feb 2012	6			
2	DFER	approve	<i>Cum Civ 22 Feb 07</i>	7			
3	DFPY	action		8			
4				9			
5				10			

SURNAME OF ACTION OFFICER AND GRADE	SYMBOL	PHONE	TYPIST'S INITIALS	SUSPENSE DATE
Dr. Leonard Kahn, AD-23	DFPY	333-8639	LK	Not Applicable
SUBJECT Clearance for Material for Public Release				DATE
USAFA-DF-PA-53				20120213

SUMMARY

1. PURPOSE. To provide security and policy review on the document at Tab 1 prior to release to the public.

2. BACKGROUND.

Authors: Leonard Kahn (sole author), Associate Professor of Philosophy, United States Air Force Academy

Title: "Strong Instrumentalism about Normative Reasons for Action"

Description: An argument for the view that motives are a necessary condition for having normative reasons for action.

Release Information: For presentation at conferences held at Louisiana State University and New Mexico State University.

Recommended Distribution Statement:

(Distribution A, Approved for public release, distribution unlimited.)

3. DISCUSSION.

4. VIEWS OF OTHERS.

5. RECOMMENDATION.

Strong Instrumentalism about Normative Reasons for Action

1. Introduction

What is the relationship between our normative reasons for action and our motives? Can one have a normative reason, say, to eat broccoli even if one does not have motivation to do so? Are matters any different if we ask about reasons for, e.g., saving a child from drowning? Can anything count in favor of taking action in the absence of motivation to do so? Instrumentalism, at least on a first approximation, answers these and similar questions as follows: An agent has a normative reason to do a particular action in a given set of circumstances only if the agent is related in the correct way to her motivate to do the action.¹

Instrumentalism occupies a paradoxical place in recent discussions of ethical theory and the philosophy of action. On the one hand, many – quite correctly – identify instrumentalism as the dominant view both in philosophy and in the social sciences.² On the other hand, instrumentalism is often dismissed as manifestly false. It has, to take only three examples, been called "a flailing theory" (Vogler, 2002: 15) which involves "a distorted conception of our standing as rational agents" (Wallace, 2006: 50) and which is "entirely unwarranted" (Thomson, 2008, 254). Yet how, one might ask, can instrumentalism be widespread among thoughtful, intelligent people and also be so patently false? What gives?

Perhaps a partial explanation is that instrumentalism is often a background feature of larger projects, even within philosophy. For example, instrumentalism is better understood as

¹ Roughly the same view is sometimes called *Humeanism* or *internalism*. Darwall (1997) parse these terms in exquisite detail. While none of these terms is ideal, I think that the term *instrumentalism* is the least problematic. Even so understood, instrumentalism is open to a number of similar but distinct interpretations. I spell out the details of the interpretation I have in mind below.

² To take just two, see Nozick (1994) and Schroeder (2007).

an assumption, rather than a theorem, in the defense of certain versions of moral relativism³ and contractarianism⁴. But let me be clear about two things. First, I do not mean that no one has ever argued for instrumentalism; that is obviously false. Indeed, the arguments of Bernard Williams (1979) have been widely discussed for decades. However, these arguments remain unpersuasive to many (to put it mildly) and are remarkably obscure as well.⁵ Second, it is not always a mistake simply to assume that a view such as instrumentalism is correct. On the contrary, every philosophical project requires at least some assumptions, and I will make plenty of my own assumptions later in this paper. That said, we should not ignore the need for instrumentalists to provide an argument for their position sooner or later - and this seems as good a time as any. So in this paper, I clarify the nature of instrumentalism and formulate what I call *the possibility argument* for this view (section 2). Then show why its first (section 3) and second (section 4) premises are credible. But *caveat emptor*. It would be too much to claim to offer anything approaching a definitive proof of instrumentalism in a brief exercise such as this. That said, I hope the reader will agree by the end of it that I have advanced debate about this issue to some extent and perhaps even offered some novel reasons for preferring one side of it.

2. Strong Instrumentalism

A few more distinctions and clarifications are in order before going any farther. First, let me let me distinguish between Strong and Weak Instrumentalism. According to what I call *Strong Instrumentalism* (SI), an agent has a normative reason to do a particular action in a

³ E.g., Harman (1975) and Wong (2006).

⁴ E.g., Gauthier (1986).

⁵ See also Williams (1985), Williams (1989), Williams (1995), and Williams (2001). Discussion of Williams' work on this subject is too vast to begin citing here.

given set of circumstances only if the agent is actually motivated to do the action, while according to *Weak Instrumentalism* (WI), an agent has a normative reason only if she is would be motivated to do the action under a particular set of conditions. The main difference between SI and WI, then, is modal: SI centers on actual motivation, and WI centers on possible motivation. Advocates of WI differ with one another about exactly what the nature of this particular set of conditions⁶, but this fact will not give us reason to pause here. For I focus in this paper entirely on SI.

It will be useful to have a fairly precise definition of *Strong Instrumentalism* with which to work:

(STRONG-INST) For all agents S and actions F, S has a normative reason to do F in circumstances C only if S has some motive M such that S's doing F in C raises the prior probability that M will be realized.

Let me also say a word or two about two of the terms used in STRONG-INST – *reason* and *motive*. To begin with, the term *reason* and its cognates are subject to many – perhaps too many – distinctions. My only concern here is with normative reasons for action, which I take to be three-part relations which hold between an agent, a possible action, and a possible state of affairs, such that the possible state of affairs counts in favor, at least to some extent, of the agent doing the possible action.⁷ We can denote this relation as follows:

⁶ See, e.g., the discussion in Finlay and Schroeder (2008, Section 1.1.2).

⁷ Compare, e.g., Scanlon (1998), Dancy (2003), and Parfit (2011).

- R(agent, possible action, possible state of affairs).

I do not assume that such a normative reason is identical with a motive or is reducible to a motive. Moreover, I do not assume that having a motive is a sufficient condition for having a normative reason. These are questions for another day. Here I am only interested in whether having a motive is a necessary condition for having a normative reason. (When speaking of a reason hereafter, I generally omit the qualifiers *normative* and *for action* for the sake of brevity.)

As for a motive, I take it to be a mental state with a mind-to-world fit, i.e., a state which moves the agent to change the world, if the world does not conform to its content⁸. A desire, of course, is a paradigmatic motive; e.g., my desire to avoid getting wet moves me to take my umbrella as I walk out the door. But desires are not unique in this way. E.g., emotions (e.g., fear of a snake) and even qualitative experiences (e.g., the pain caused by burning one's hand on an open flame) can also serve as motives. Perhaps other mental states can motivate as well. Can a belief motivate? I have elsewhere argued that it cannot, but that point is neither here nor there for my present purposes, so I leave this question open here.⁹ I have not, to be sure, offered anything like definitions of *reason* and *motive*, and even the rather bland claims I have made about reasons and desires are not universally accepted. Nevertheless, they are sturdy enough to support the work I do in this paper.

But is Strong Instrumentalism correct? Yes, according to what I call the *possibility argument*:

⁸ See, e.g., Smith (1994) and Searle (2002).

⁹ See my [information suppressed to retain anonymity of the current author]. For the contrary position, see, e.g., Nagel (1970), McDowell (1979), McNaughton (1987), Schueler (1997), and Dancy (2000).

1. S has a reason to do F in C only if she can do F in C.
2. S can do F in C only if she is motivated to do F in C.
3. S is motivated to do F in C only if S has some motive M such that S's doing F in C raises the prior probability that M will be realized.

--

4. Therefore, S has a reason to do F in C only if S has some motive M such that S's doing F in C raises the prior probability that M will be realized.¹⁰

Proposition 4 is, of course, simply a restatement of STRONG-INST. I take Premise 3 to be an analytic truth, so in the remainder of the paper I make the case for Premises 1 and 2.

3. The First Premise

Famously, the claim that *ought* implies *can* turns on how we understand the term *can*.¹¹ In a similar manner, the claim under consideration here - that *has a reason* implies *can* - does so as well. This is not the place to spell out every nuance of the relevant concepts. On the contrary, I need only to show that there is a single sense of the term *can* according to which both Premise 1 and Premise 2 are true.¹²

But what is this sense of *can*? One way to try to answer this question would be to

¹⁰ Assume that Propositions 1 through 4 quantify over all agents, S, and all actions, F.

¹¹ Kant ([1781] 1999). The literature is enormous, but for excellent recent work see, among others, Yaffe (1999), Saka (2000), Mellema (2001), Mason (2003), Struemer (2003; 2008; Forthcoming), Kramer (2005), Howard-Snyder (2006), Väyrynen (2006), Vranas (2007), Copp (2008), Littlejohn (2009), and Heuer (2010).

¹² Let me stress the difference between the possibility argument and other recent attempts to use something similar to the idea that 'ought' implies 'can' to defend something similar to instrumentalism. Anomaly (2008) attempts to make the case for Williams' version of WI via the lack of a "sound deliberative route" from the agents motives to her actions. While this is a project worth pursuing, it is not my project.

specify a familiar modal sense of *can* such as the *can* of metaphysical possibility (according to which there is some possible world *W* in which *S* does *F*) or the *can* of nomological possibility (according to which there is some possible world *W* with the same laws of nature as our own in which *S* does *F*).¹³ But these senses of *can* are largely too crude for the task at hand.

So let me sketch the contours of the sense of the term *can* which I have in mind by means of a thought experiment which I shall call:

Mary and the Mountain Lion: Suppose that while hiking Mary stumbles across a mountain lion. Mary is alone, unarmed, and unable to outrun this creature. Luckily for her, Mary remembers reading about how to respond to such a creature. "Raise your arms," her reading instructed, "and make yourself appear bigger than you are." Mary certainly has a reason to make herself appear bigger than she is, perhaps by raising her arms over her head. For this will make the mountain lion less likely to attack her.

Now consider a first variation of this thought experiment: Though Mary certainly has a reason to make herself appear bigger than she is, she does not have reason to act in such a way that she is bigger than herself. Why? For starters, notice the obvious point: nothing can be bigger than itself, and there is, therefore, no possible state of affairs identical with *X*'s being bigger than itself. Recall that in Section 2 of this paper I observed that a reason for action is a

¹³ See, e.g., Salmon (1989) and Chalmers (1997).

relation between an agent, a possible action, and a possible state of affairs, such that the possible state of affairs counts in favor of the agent doing the possible action. But since there is no possible state of affairs identical with X's being bigger than itself there is no state of affairs that can play the role of the third relatum in

- R(agent, possible action, possible state of affairs).

So there are no reasons for action which count in favor of becoming bigger than yourself because there are no possible states of affairs that can satisfy the reason relation. Here, then, is a clear sense of can in which has a reason implies can, that is to say, a clear sense of can in which Premise 1 is true.

Turn now to a second variation which illustrates this same sense of *can*. I have already noted that Mary might try to appear bigger than she actually is by raising her arms over her head. But there might be better ways to scare off a mountain lion. Here's one: Mary would be more likely to appear large if she were to stand on stilts. So it is unfortunate that Mary does not have stilts with her. Now, given that Mary does not have stilts, it is not possible for her to stand on them. That is to say, there is no possible world in which Mary both does not have stilts with her and does stand on them. Since there is no such possible world, there is no possible action which Mary undertakes that is properly described as standing on stilts. So, in this case, Mary does not have a reason to scare off the mountain lion by standing on stilts because there is no possible action in which she does so. In this variation, the problem is not that there is no possible state of affairs in which Mary looks bigger than she is. The problem,

rather, is that given her lack of stilts, she cannot stand on them. So there is no possible action she can take in which she can stand on them and, therefore, no reason for her to do so.

To be sure, what I have just provided is only two illustrations of this sense of *can*, but it is a simple matter to generalize on the basis of this illustration, and that is all that is required (or possible!) in a short paper like this one.

4. The Second Premise

Now the question is whether Premise 2 is true given that we understand *can* in the same sense. Here is a brief argument for answering this question in the affirmative:

1. S has a reason to do F in C only if S can do F in C.
 - 1a. S can do F in C only if, in C, she has what she needs to do F.
 - 1b. S, in C, has what she needs to do F only if she is motivated to do F.
2. Therefore, S can do F in C only if she is motivated to do F in C.¹⁴

Of course, I have already explicitly argued that Premise 1 is correct in the last section. However, a discerning reader will note that I also offered some intuitive considerations that favor Premise 1a as well. In the second variation, it is Mary's lack of what she needs to stand on stilts that explains her lack of reason to stand on stilts. That is to say, it is because it is false that Mary has what she needs to stand on stilts that it is also false that she can stand of

¹⁴ Assume that Propositions 1, 1a, 1b, and 2 quantify over all agents, S, and all actions, F.

stilts. Moreover, it is because it is false that Mary can stand on stilts that she has reason to do so. Note that I do not claim more than Premise 1a here. More specifically, I do not claim that one can do whatever one needs to do the thing in question. Perhaps there are other possible defeating conditions. That is not an issue that I plan on settling in this paper.

Let us turn now to Premise 1b. In order for something, F, to count as S's action, F must be something that S does, not merely something that happens to S. It may well be that several things need to be true in order for something to count as S's action, but one of them surely must be that S has a motive for doing F. Here is why I think that is the case: Either S's action F has a cause or it does not. If F does not have a cause, then, trivially, it does not have S as a cause and so cannot be correctly attributed to S. If F does have a cause, then, either some property of S is one of the causes, or it is not. If some property of S is not one of the causes of F, then, once again F cannot be correctly attributed to S. If some property of S is one of the causes of F, then it is at least possible that F can be correctly attributed to S. So it is at least possible that F can be correctly attributed to S only if one of the causes of S doing F is some property of S. But what kind of property could this be? The answer is that it must both be a feature of a motivational system and be S's not someone else's. It must be a feature of a motivational system since that is all that serve as a cause for an agent, and it must be S's not someone else's because this is the only thing that could make it S's action rather than P's or Q's. And the only feature that could play both of these roles is one of S's motives. So among the things that S must have in order for F to be her action is the motive to do F. And S's motive for doing F motivates her to do F. Therefore, S, in C, has what she needs to do F only if she is motivated to do F.

So much for the argument. Let me now remove some possible sources of confusion.

First, there are surely some irrelevant senses of *can* in which it is true that S can do F even if she is not now motivated to do F. S might be physically forced to do an action, e.g., hit herself in the face. That is to say S might have her hand formed into a fist and have the fist smashed into her face. But this would not be S's action in anything other than the imagination of a school-yard bully. So too S might be coerced into doing F, say, by having her life threatened if she did not. Yet in this case S does have a motive to do the action (her desire to avoid death) and almost certainly gains a reason to do so as well.

Second, I do not, of course, claim that if one is not now motivated to do F in C one cannot become motivated to do so. Perhaps one can. I am not now motivated to write a novel, but I might become motivated to do so next year. So the proposition that I am not motivated to do F in C at time t1 entails that I do not at t1 have a reason to do F in C, but it does not entail that I cannot come to have such a reason later at, say, t2. And if I come to be motivated to do F in C, I may come to have a reason to do F in C. But the proposition that I can have a reason in the future does not commit me to saying I do have a reason now.

Third, I also do not claim that one cannot have a reason to be motivated to do something that one is not now motivated to do. This paper is about reasons for actions, not reasons for mental states, and I make no claims here about the latter. But it does not follow from the fact (if it is a fact) that I have a reason to be motivated to do F that I have a reason to do F. For instance, one might be offered a huge reward for wanting to drink a toxin but nothing for actually doing so.¹⁵ In this case one has a reason to want to do F but no reason to do F.

I have argued, all too quickly, that S has a reason to do F in C only if S has some desire D such that S's doing F in C raises the prior probability that D will be satisfied. Most of

¹⁵ See, e.g., Kavka, (1983) and Parfit (1984).

the argument has consisted of showing the plausibility of the following two claims: S has a reason to do F in C only if she can do F in C, and S can do F in C only if she is motivated to do F in C. Far more work remains to be done, but I will be quite happy if I have managed to do so.

References

- Anomaly, Johnny 2008. "Internal Reasons and the Ought-Implies-Can Principle," *Philosophical Forum* 39(4), 469-483.
- Chalmers, David 1997. *The Conscious Mind*. Oxford: Oxford University Press.
- Copp, David 2008. "'Ought' Implies 'Can' and the Derivation of the Principle of Alternate Possibilities," *Analysis* 68(297), 67-75.
- Dancy, Jonathan 2003. *Ethics without Principles*. Oxford: Oxford University Press.
- Dancy, Jonathan 2000. *Practical Reality*. Oxford: Oxford University Press.
- Darwall, Stephen 1997. "Reasons, Motives, and the Demands of Morality: An Introduction," in *Moral Discourse and Practice*. Edited by Stephen Darwall, Allan Gibbard, and Peter Railton. Oxford: Oxford University Press.
- Gauthier, David 1986. *Morals by Agreement*. Oxford: Oxford University Press.
- Harman, Gilbert 1975. "Moral Relativism Defended" in *Explaining Value and Other Essays in Moral Philosophy*. Oxford: Oxford University Press.
- Heuer, Ulrike 2010. "Reasons and Impossibility," *Philosophical Studies* 147(2).
- Howard-Snyder, Frances 2006. "'Cannot' Implies 'Not Ought,'" *Philosophical Studies* 130(2), 233 - 246.
- Kant, Immanuel [1781] 1999. *Critique of Pure Reason*. Edited and Translated by Paul Guyer and Allan Wood. Cambridge, UK: Cambridge University Press.
- Kavka, Gregory 1983. "The Toxin Puzzle," *Analysis* 43, 33-36.
- Kramer, Matthew H. 2005. "Moral Rights and the Limits of the Ought-Implies-Can Principle: Why Impeccable Precautions Are No Excuse," *Inquiry* 48(4), 307 - 355.
- Littlejohn, Clayton 2009. "'Ought', 'Can', and Practical Reasons," *American Philosophical Quarterly* 46(4), 363-73.
- McDowell, John 1979. "Virtue and Reason," *Monist* 42, 331-350.
- McNaughton, David 1987. *Moral Vision*. Oxford: Blackwell.
- Mason, Elinor 2003. "Consequentialism and the Ought Implies Can Principle," *American Philosophical Quarterly* 40(4), 319-331.
- Mellema, Gregory 2001. "Praise, Blame, and the Ought Implies Can Principle," *Philosophia* 28(1-4), 425-436.
- Nagel, Thomas 1970. *The Possibility of Altruism*. Princeton, New Jersey: Princeton University Press.
- Nozick, Robert 1994. *The Nature of Rationality*. Princeton, New Jersey: Princeton University Press.
- Parfit, Derek 2011. *On What Matters*, Volume 1. Oxford: Oxford University Press.

- Parfit, Derek 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- Saka, Paul 2000. "Ought Does Not Imply Can," *American Philosophical Quarterly* 37(2), 93 - 105.
- Salmon, Nathan 1989. "The Logic of What Might Have Been" in *Metaphysics, Mathematics, and Meaning: Philosophical Papers Volume 1*. Oxford: Oxford University Press.
- Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
- Schroeder, Mark 2007. *Slaves of the Passions*. Oxford: Oxford University Press.
- Schueler, Fred 1997. *Desire*. Cambridge, Massachusetts: MIT Press.
- Searle, John R. 2002. *Rationality in Action*. Cambridge, MA: MIT Press.
- Smith, Michael 1994. *The Moral Problem*. Oxford: Blackwell.
- Streumer, Bart (Forthcoming). "Reasons, Impossibility and Efficient Steps: Reply to Heuer," *Philosophical Studies*.
- Streumer, Bart 2007. "Reasons and Impossibility," *Philosophical Studies* 136(3), 351-384.
- Streumer, Bart 2003. "Does 'Ought' Conversationally Implicate 'Can'?", *European Journal of Philosophy* 11(2), 219-228.
- Thomson, Judith Jarvis 2008. *Normativity*. New York, NY: Open Court.
- Väyrynen, Pekka 2006. "Ethical Theories and Moral Guidance," *Utilitas* 18(3), 291-309.
- Vogler, Candace 2002. *Reasonably Vicious*. Cambridge, MA: Harvard University Press.
- Vranas, Peter B. M. 2007. "I Ought, Therefore I Can," *Philosophical Studies* 136(2), 167 - 216.
- Wallace, R. Jay 2006. *Normativity and the Will*. Cambridge, MA: Harvard University Press.
- Williams, Bernard 1979. "Internal and External Reasons" in *Moral Luck and Other Essays*. Cambridge, UK: Cambridge University Press.
- Williams, Bernard 1985. *Ethics and the Limits of Philosophy*. Cambridge MA: Harvard University Press.
- Williams, Bernard 1989. "Internal Reasons and the Obscurity of Blame" in *Making Sense of Humanity and Other Philosophical Papers*. Cambridge, UK: Cambridge University Press.
- Williams, Bernard 1995. "Replies" in J.E.J. Altham and R. Harrison (eds.) *World, Mind and Ethics: Essays on the Ethical Philosophy of Bernard Williams*. Cambridge, UK: Cambridge University Press.
- Williams, Bernard 2001. "Postscript: Some Further Notes on Internal and External Reasons" in E. Millgram (ed.) *Varieties of Practical Reasoning*. Cambridge MA: MIT Press.
- Yaffe, Gideon 1999. "'Ought' Implies 'Can' and the Principle of Alternate Possibilities,"

Analysis 59(3), 218-222.